

自動チューニング機能付き 数値計算ライブラリXabclib

片桐孝洋

東京大学情報基盤センター

共同研究者

大島聡史（東京大学情報基盤センター）、黒田久泰（愛媛大学/東京大学情報基盤センター）、伊藤祥司、中島研吾（東京大学情報基盤センター）、櫻井隆雄、直野健（（株）日立製作所中央研究所）、猪貝光祥（（株）日立超LSIシステムズ）

T2Kシンポジウム2010@東大

日時：2010年7月12日（月）13:00～17:20

場所：東京大学武田先端知ビル 5階武田ホール

第2部 eScienceプロジェクト成果報告

14:00～14:20

発表の流れ

1. 背景
2. OpenATLib と Xabclib (β版)
 - OpenATLib:
汎用自動チューニングインターフェース
 - Xabclib:
OpenATLibを用いた自動チューニング機能付き数値計算ライブラリ
3. 数値計算ポリシー機能
 - 実行速度
 - メモリ量
 - 演算精度 の最適化が可能
4. おわりに

計算機環境の複雑化

- **マルチコアの浸透**
 - 非均質メモリアクセス (ccNUMA)
 - 多階層化されたキャッシュ構造
 - L1・L2は局所、L3は共有
 - チップ内のコア数の増大
 - ハイエンド：32コア～、ローエンド：8コア～
- **並列実行モデルの変化**
 - ピュアMPI実行
 - 低～中並列実行 (～1,000コア) 時
 - **ハイブリッドMPI** (OpenMP + MPI) 実行
 - 超並列実行 (10,000+コア実行) 時

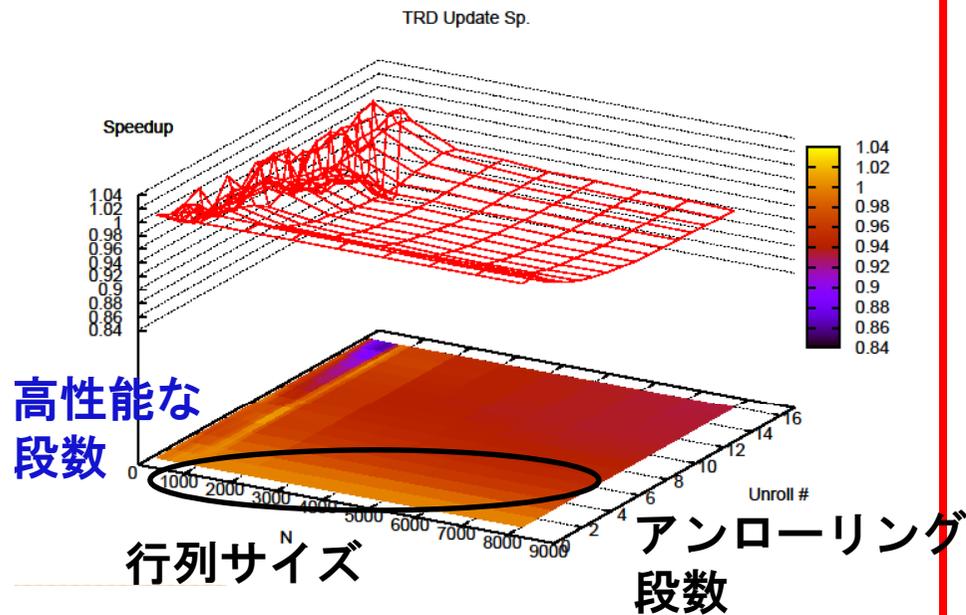
実行時自動チューニングの必要性

- **現実の性能データはモデル化できない**
 - **予測できない「スパイク」がある**
 - **入力データ依存性がある**
 - 実行時に定まる、疎行列形状 や 数値特性
 - 最適な 前処理方式 や 利用アルゴリズムが 実行時まで不明
 - **OSのリソースマネージメント依存**
 - 割り当てられる ノード や ソケットが**実行時まで不明**
 - ローカルメモリがつかえるか
 - マスター・スレーブ形態で、理想的にコア割り当てされるか
 - 割り当てノード群に対する、**物理ネットワーク形状（通信性能）が実行時まで不明！**
- **コンパイラ（だけ）では手に負えない**
 - **手動チューニングはコスト高**
 - **手動でできないこともある**
 - 予測できない極端な性能差が不確定におこる場合
 - **実用的観点から性能自動チューニング技術へ期待**

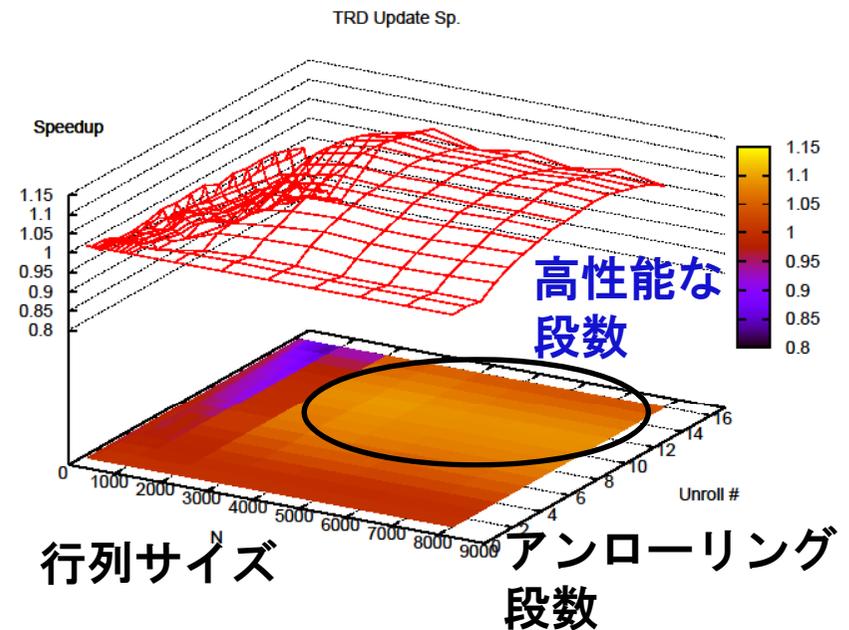
マルチコアの実行形態の違いによる 性能挙動の変化 (T2Kオープンスパコン)

対称固有値ソルバの三重対角化 (行列更新部分)

4コア (4コア / 1ソケット実行)



4コア (1コア / 1ソケット実行)



ソケットごとの共有L3キャッシュの影響

→ (動的な) 実行コア数により最適実装が変わる!

e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発
 高生産・高性能計算機環境実現のためのシステムソフトウェアの研究開発
 「シームレス高生産・高性能プログラミング環境」高性能高可搬性ライブラリ
 (平成20年度～平成23年度) : 代表 東京大学 石川裕 教授

● 問題

- 最適化が職人芸に依存、生産性がない、可搬性がない、煩雑な処理、製作コストがかかる
- 理論的に範囲外のパラメタ指定による収束失敗

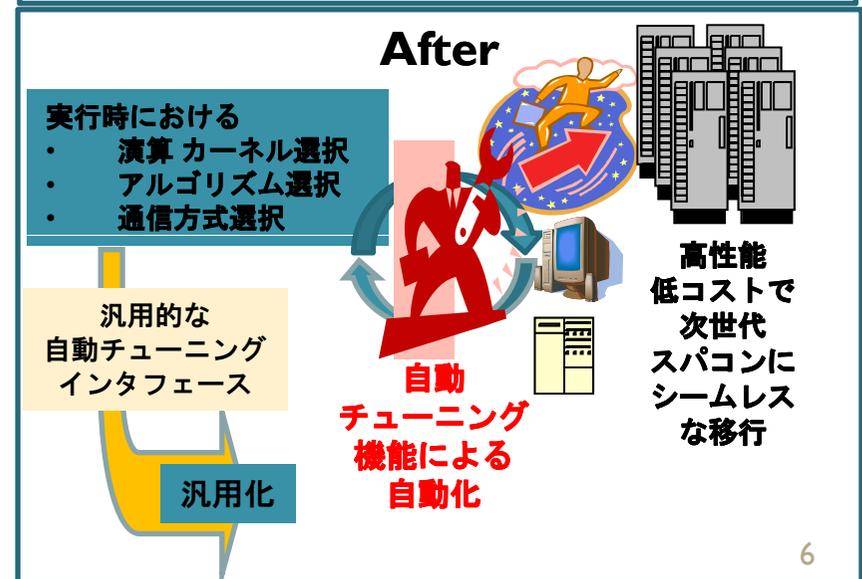
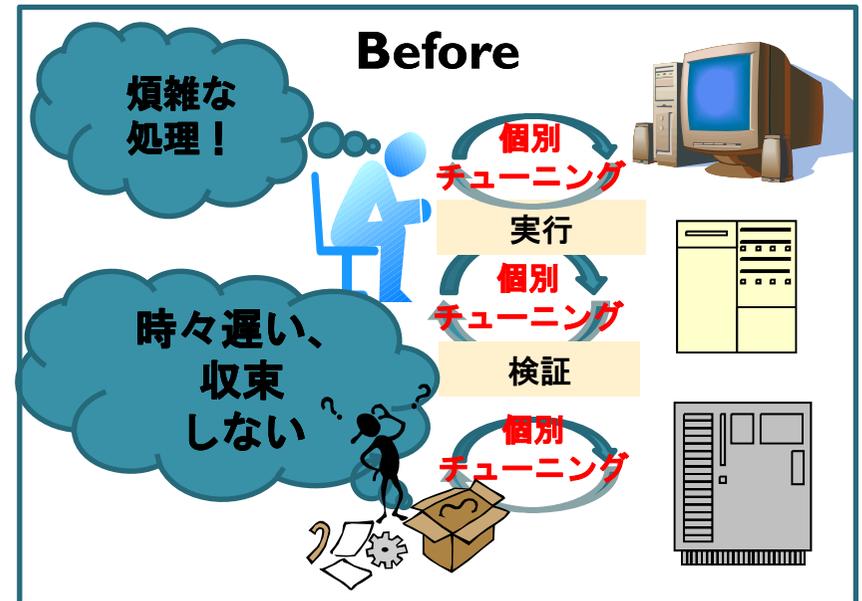
● 目標

- 高生産性／性能移植性をもつ数値計算ライブラリの提供

● 疎行列の非ゼロ成分の分布を考慮し

実行時における以下の自動チューニング機能を提供

1. 演算カーネル選択
2. 数値アルゴリズム選択
3. 並列実装方式選択
4. 汎用的なAPI (OpenATLib)



OpenATLibの位置づけ

- 動的選択する自動チューニングインターフェースを（疎行列を扱う）ATライブラリ開発者やユーザに提供
- 以下の性能パラメタ最適化機能を提供
 1. アルゴリズム選択処理
 - Krylov部分空間のリスタート周期選択
 - 直交化方式選択
 2. 実装方式選択
 - 疎行列 - ベクトル積 (SpMxV)
 - 負荷バランス改良、ベクトル計算機向け、並列化向き
 3. 計算機資源選択
 - 実行時のメモリ量・コア数などを考慮したSpMxVとリスタート周期
- ユーザポリシー設定機能
 - 実行速度、メモリ量、演算精度

Xabclib (β版) の特徴

- **汎用的自動チューニングAPIの提供**
 - 自動チューニング機能の再利用性を考慮した、汎用的APIのOpenATLibを提供
- **世界に先駆けた先進的AT機能の提供**
 1. **自動チューニングポリシー設定機能**
 - 実行速度、メモリ量、演算精度の方針設定機能
 2. **マルチコア向きの負荷バランスの改善が可能な疎行列演算方式**
 - 非零要素均等化方式
 - Branchless Segmented Scan (BSS) 方式
 3. **直交化アルゴリズム動的選択機能**

OpenATLib (β版)で提供される関数

関数名	機能の説明
OpenATI_DAFRT	Krylov部分空間のリスタート周期を増加するか判定
OpenATI_DSRMV	CRS形式において、倍精度演算 対称行列用の疎行列ベクトル積で、最もよい実装方式の選択
OpenATI_DURMV	CRS形式において、倍精度演算 非対称行列用の疎行列ベクトル積で、最もよい実装方式の選択
OpenATI_DSRMV_Setup	OpenATI_DSRMVのための初期データ設定処理
OpenATI_DURMV_Setup	OpenATI_DURMVのための初期データ設定処理
OpenATI_DAFGS	Gram-Schmidt直交化処理において、4種類の実装から最もよい実装方式の選択
OpenATI_BLDATA	デフォルトデータ設定処理 (Fortran言語のBlock data format)
OpenATI_LINEAR SOLVE	数値計算ポリシーを適用する連立一次方程式ソルバ用 メタ・インタフェース
OpenATI_EIGENSOLVE	数値計算ポリシーを適用する固有値ソルバ用 メタ・インタフェース

疎行列-ベクトル積の実装方式

- OpenATI_DSRMV (対称行列用 API)
 - S1) 行分割方式 (RDM)
 - S2) 非零要素均等化方式 (NNZ)(負荷バランス化RDM)
 - S3) 並列リダクション時に零演算を除去したADM
- OpenATI_DURMV (非対称行列用 API)
 - U1)行分割方式 (RDM)
 - U2)非零要素均等化方式 (NNZ)(負荷バランス化RDM)
 - U3) Branchless Segmented Scan(BSS) [1]
(スカラマシン用)
 - U4) オリジナルSegmented Scan [2]
(ベクトルマシン用)

[1] Takao Sakurai, et.al. SIAM PP2010, MS6 (2010)

[2] Guy E. Blelloch, Michael A. Heroux, and Marco Zagha: Segmented Operations for Sparse Matrix Computation on Vector Multiprocessors, Technical Report of CMU-CS-93-173 (August 1993)

行分割方式(RDM)と非零要素均等化方式(NNZ) (対称と非対称行列双方)

* : 非零要素

割り当て
要素数

割り当て
要素数

スレッド1	*						9
スレッド2		*					
スレッド3			*				
スレッド4			*	*	*		
				*	*	*	
	*		*			*	

	*						3
		*					
			*				
			*	*	*		
				*	*	*	
	*		*			*	

行分割方式 (RDM)
(SW: SI, UI)

非零要素均等化方式 (NNZ)
(SW: S2, S3, U2)

それぞれのスレッドでの割り当て
行数を均等化するように分割。

それぞれのスレッドで割り当てられる
非零要素数を均等化するように
割り当て行数を決定。

Segmented Scan と BSS (非対称行列のみ)

input matrix

1	0	0	2	0	0	0	0
0	3	0	0	0	4	5	6
7	8	0	9	0	10	0	0
11	0	12	13	0	0	14	0
0	0	0	0	15	0	16	17
18	19	20	21	22	23	24	25
0	26	0	27	0	0	0	0
0	0	0	0	0	0	28	0
29	0	0	0	0	0	30	0

row pointer (IRP)

1	3	7	11	15	18	26	28	29	31
---	---	---	----	----	----	----	----	----	----

非零要素 6 個ごとに分割。
行単位で分割ではない！

segment-vector #3

Branchless Segmented Scan (BSS)

MFLAG	JFSTART
1	0
3	2
7	4
11	7
13	8
15	12
18	
19	
25	
26	
28	
29	
31	

JFSTARTは
各セグメントの
最終要素の位置
を保持

これにより IF-文が
プログラムで不要に！

アクセス方向

○ 各行の
最初の要素

オリジナルな
Segmented Scan

FLAG

T	T	F	F	F
F	F	F	F	T
T	F	T	F	F
F	F	F	F	T
F	T	F	F	T
F	F	T	F	F

JL: Segment Length

“T” が各行の
最初の要素
を意味する

アクセス方向

数値計算ポリシー機能

- 右のような「設定ファイル」を用意
- 実行例（報告ファイルの中身）
 - OPENATI_POLICY_REPORT.txt

```
POLICY = ACCURACY
RESIDUAL = 1.0D-10
CPU = 16
PRECONDITIONER = ILU0
MAXMEMORY = 1.0
MAXTIME = 500.0
```

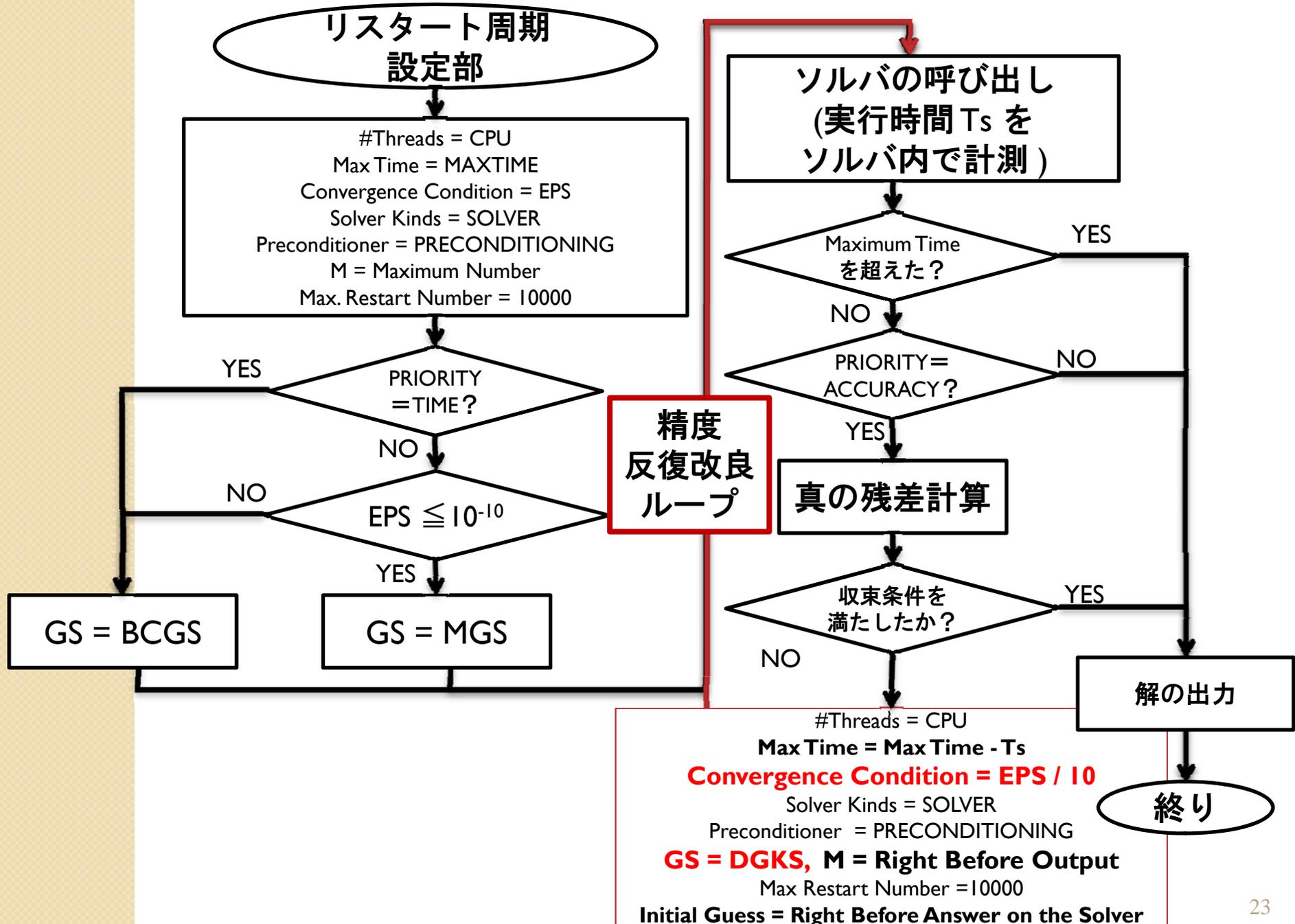
```
*****
**** OpenATI LINEAR SOLVER POLICY REPORT ****
****                2010.0114 11:30                **** ←report date / time
*****

[Environment variables]                               ↓ input parameters
OPENATI_DEBUG =
OPENATI_POLICY = ./input_policy.dat

[Policy Definitions]
POLICY = ACCURACY
SMPs = 16
SOLVER = XABCLIB GMRES
PRECONDITIONER = ILU0
REQUIREMENT WORKING MEMORY = 16.000000000000000
<<< Upper Bound 16GBYTE >>>
REQUIREMENT RESIDUAL = 1.0000000000000000E-008
REQUIREMENT MAX.TIME = 500.000000000000000
MAX. SUBSPACE SIZE = 14214
RUNTIME MEMORY USE = 3.24 [GBYTE]
KRYLOV SUBSPACE EXPAND AT = 1 ,MATVEC AT = 1
Initial Gram-Schmidt Strategy = BCGS

===== OPENATI LINEARSOLVE SUCCESSFULLY ENDED ===== ↓ successfully exit
[OPENATI LINEARSOLVE RESULT]                            ↓ result report
MATRIX DATA : N= 14214 NNZ= 259688
FASTEST MATVEC NO. = 11 ←fastest OpenATI_DURMV case
FINAL KRYLOV SUBSPACE SIZE = 42 ←Msize for convergence
FINAL Gram-Schmidt Strategy = DGKS
2-Norm of RHS = 25.2388589282479 ←initial norm of RHS
NUMBER OF RETRYED GMRES = 6 ←retried iterations
TOTAL RESTARTS of GMRES = 197
RESIDUAL NORM = 3.005885687924543E-010
SET-UP TIME = 1.126790046691895E-002 [SEC]
SOLVER TIME = 1.3203270435333 [SEC]
TOTAL TIME = 1.3315949440024 [SEC]
```

演算精度ポリシーのパラメタ設定戦略





性能評価

計算機環境： T2Kオーpensパソコン (東大版)

CPU	Quad-Core AMD Opteron(TM) Processor 8356 2.3GHz ,16 core/node
L1 Cache Size	Data: 64 Kbyte/core, Instruction: 64 Kbyte/core
L2 Cache Size	512 Kbyte/core
L3 Cache Size	2 MByte/4core
主記憶	32GByte (8GBytes/Socket)
OS	Red Hat Enterprise Linux 5
コンパイラ	Intel Fortran Compiler Professional Version 11.0
コンパイラ オプション	-O3 -m64 -openmp -mcmmodel=medium

数値実験の条件

●OpenATI_EIGENSOLVE（固有値ソルバ）

収束判定値	1.0E-08
計算する固有値数	10
最大実行時間	1000 [秒]

●OpenATI_LINEARsolve（連立一次方程式ソルバ）

収束判定値	1.0E-08
右辺ベクトルの値	全ての要素が 1
初期ベクトルの値	全ての要素が 0
前処理方式	ILU(0)
最大実行時間	1000 [秒]

- 従来法のデフォルトのリスタート周期
30 : PETScのデフォルト値と同じ

OpenATI_EIGENSOLVEのテスト行列

- University Florida Sparse Matrix Collection (21種、対称行列)

Matrix	N	NNZ	Field
vibrobox	12328	177578	Acoustics
Lin	256000	1011200	Chemistry
cfdl	70656	949510	Fluid Dynamics
cfdl2	123440	1605669	
gyro	17361	519260	Model Reduction
t3dl	20360	265113	
c-71	76638	468096	Optimization
c-73	169442	724348	
Si5H12	19896	379247	Structural
SiO	33401	675528	
dawson5	51537	531157	

Matrix	N	NNZ	Field
H2O	67024	1141880	Structural
F2	71505	2682895	
oilpan	73752	1835470	
shipsec1	140874	3977139	
bmw7st_1	141347	3740507	
SiO2	155331	5719417	
shipsec5	179860	5146478	
Si41Ge41H72	185639	7598452	
bmw3_2	227362	5757996	
Ga41As41H72	268096	9378286	

OpenATI_LINEAR_SOLVEのテスト行列

- University Florida Sparse Matrix Collection
(22種、非対称行列)

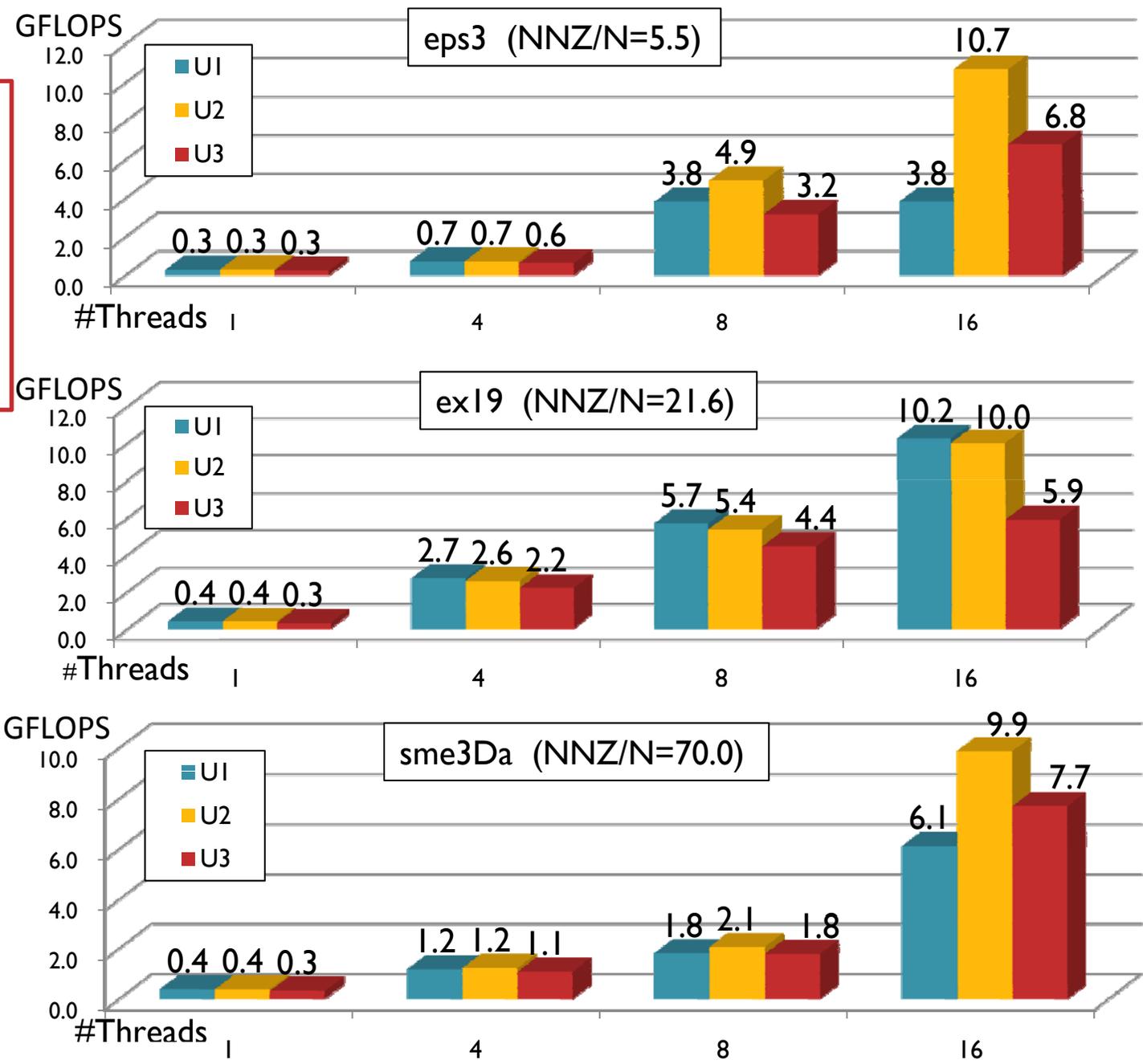
Matrix	N	NNZ	Field
chipcool0	20082	281150	2D/3D
chem_master1	40401	201201	
torso1	116158	8516500	
torso2	115067	1033473	
torso3	259156	4429042	
memplus	17758	126150	Electric circuit
ex19	12005	259879	Fluid dynamics
poisson3Da	13514	352762	
poisson3Db	85623	2374949	
airfoil_2d	14214	259688	
viscoplastic2	32769	381326	Materials

Matrix	N	NNZ	Field
xenon1	48600	1181120	Materials
xenon2	157464	3866688	
wang3	26064	177168	Semiconductor device
wang4	26068	177196	
ecl32	51993	380415	
sme3Da	12504	874887	Structural
sme3Db	29067	2081063	
sme3Dc	42930	3148656	
epb1	14734	95053	Thermal
epb2	25228	175027	
epb3	84617	463625	

典型的な非対称行列の疎行列-ベクトル積性能

U1: 行分割
U2: 非零要素均等化
U3: BSS

U2: 非零要素均等化方式が多くの場合、よい性能を達成。

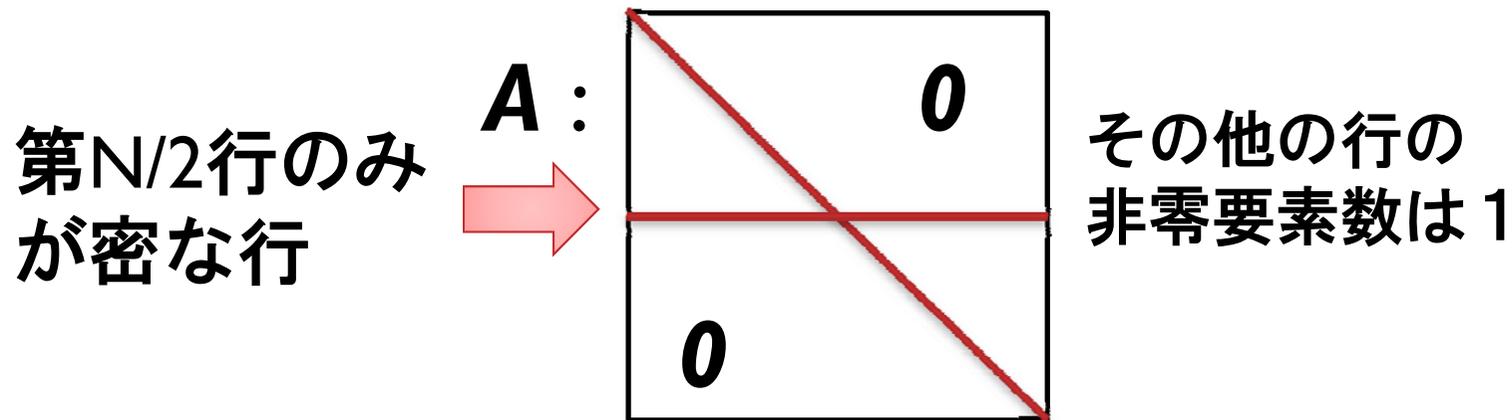


行分割方式で並列化できない例

- 以下の行列

$$A = (a_{i,j}), (i, j = 1, 2, \dots, N)$$

$$\begin{cases} \text{if } ((i = j) \text{ and } (i \neq N / 2)) \text{ then } \# \text{ non - zero} = 1. \\ \text{if } (i = N / 2) \text{ then } \# \text{ non - zero} = N. \\ \text{The elements are generated by rand } (). \end{cases}$$

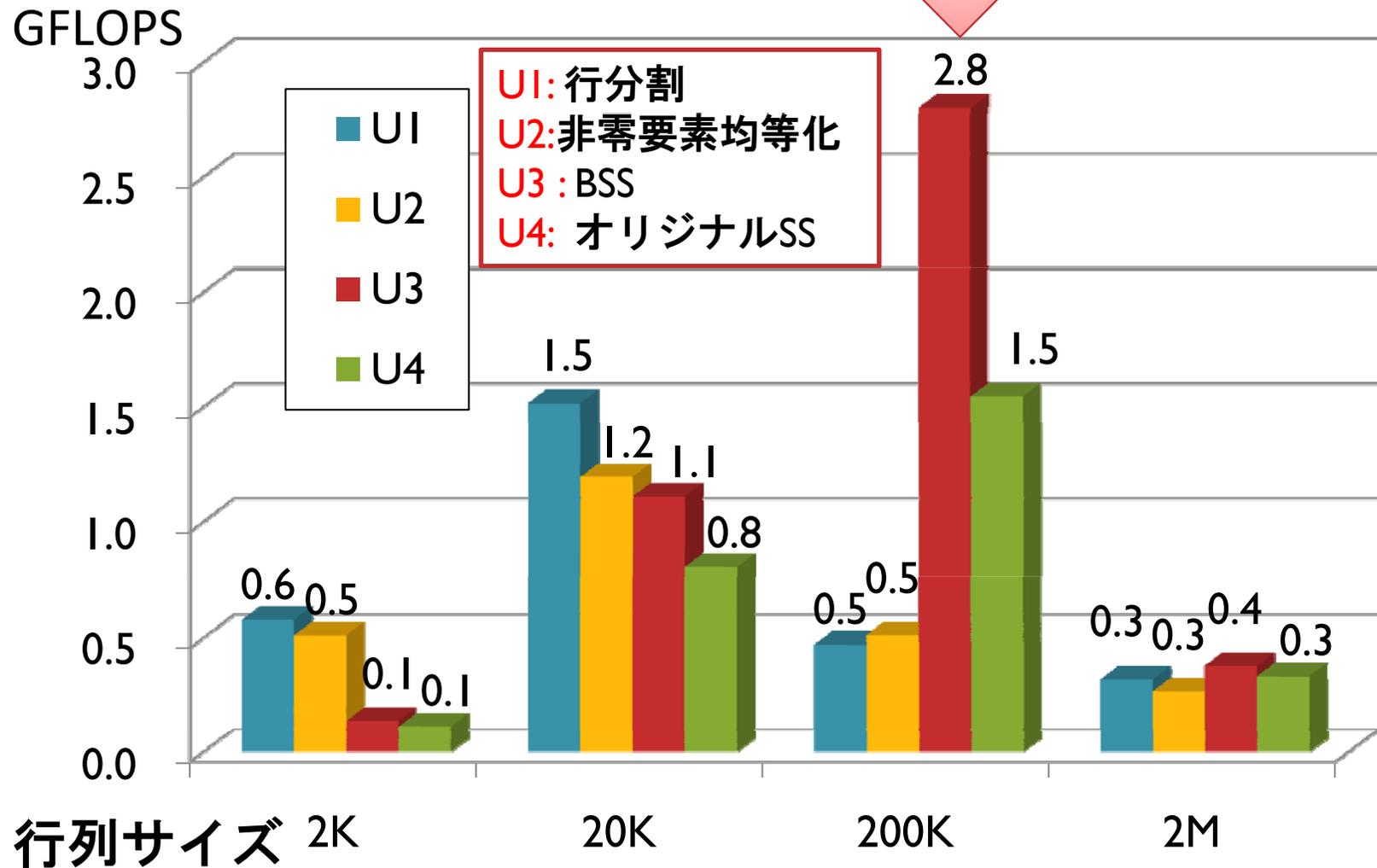


- 行分割方式では、第N/2行の計算について並列性が抽出できない

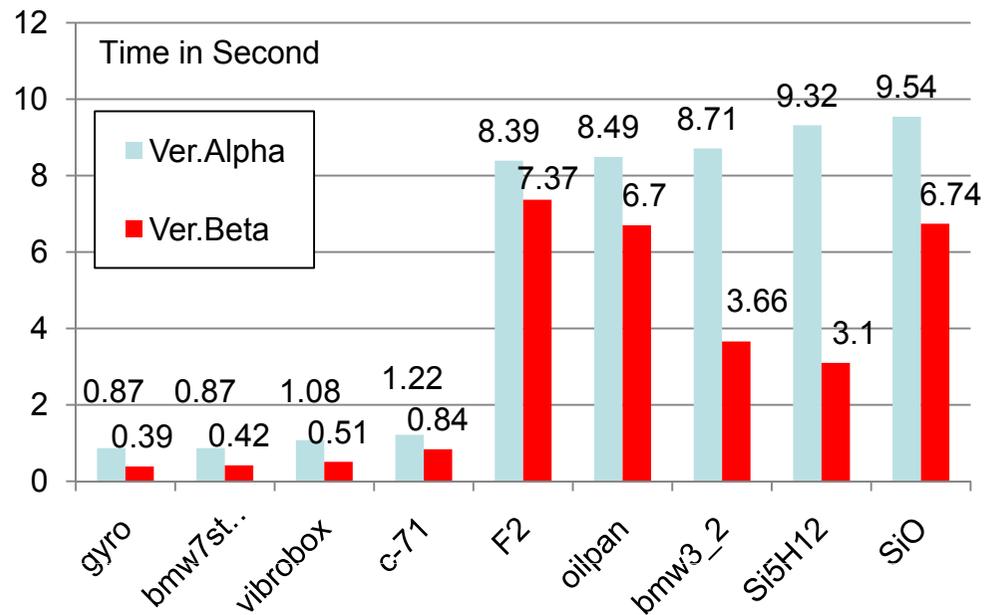
OpenATI_DURMV(β版)の性能

- #Cores = 16
- JL = 128 (Constant)

BSSで約5.6倍の速度向上



精度ポリシーの効果(OpenATI_EIGENSOLVE)



α版において、以下の行列は
要求精度を満たさず

●T3dl:

相対残差: 2.24E-01

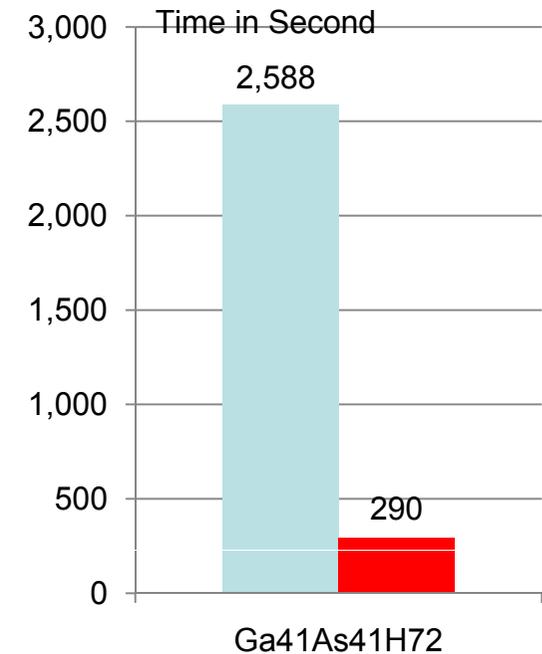
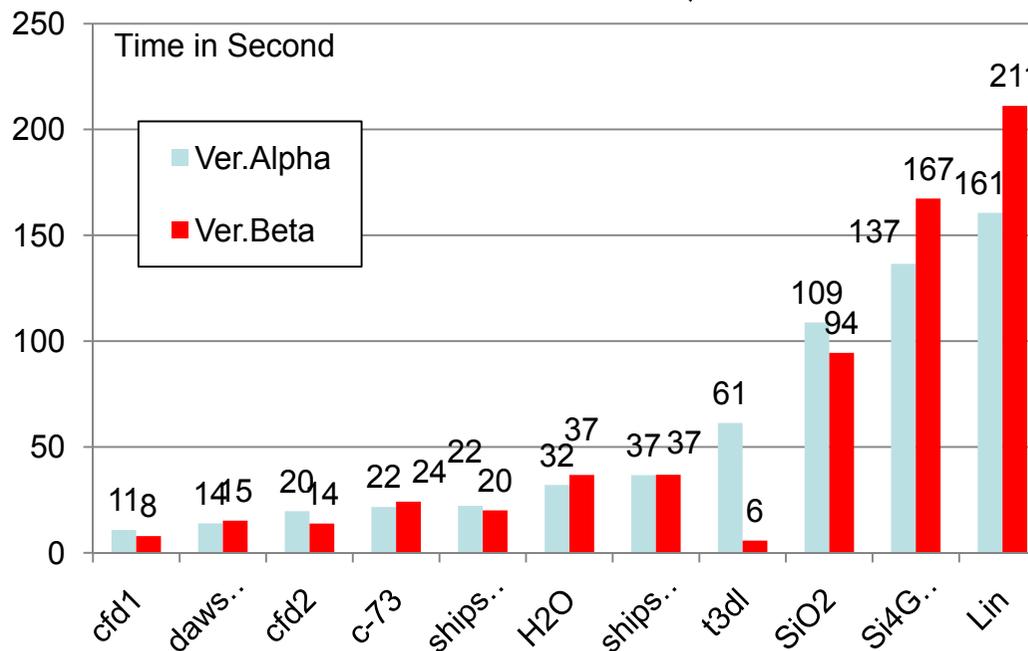
直交性: 1.27E-02

●Ga4IAs4IH72

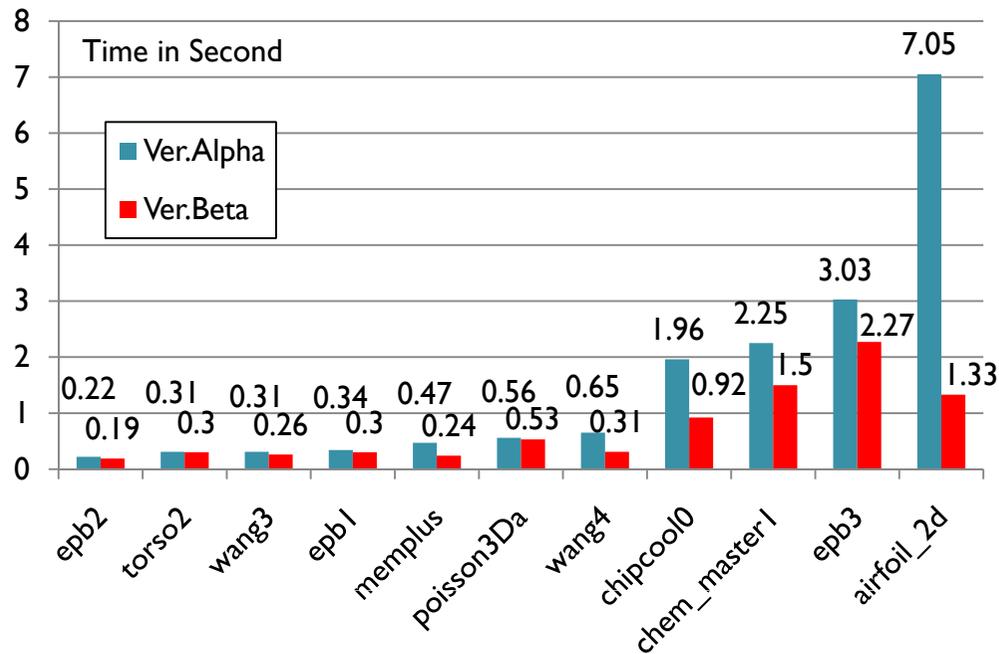
相対残差: 2.52E-08

直交性: 4.81E-09

平均速度向上: 2.24倍
最大速度向上: 8.92倍
(Ga4IAs4IH72)



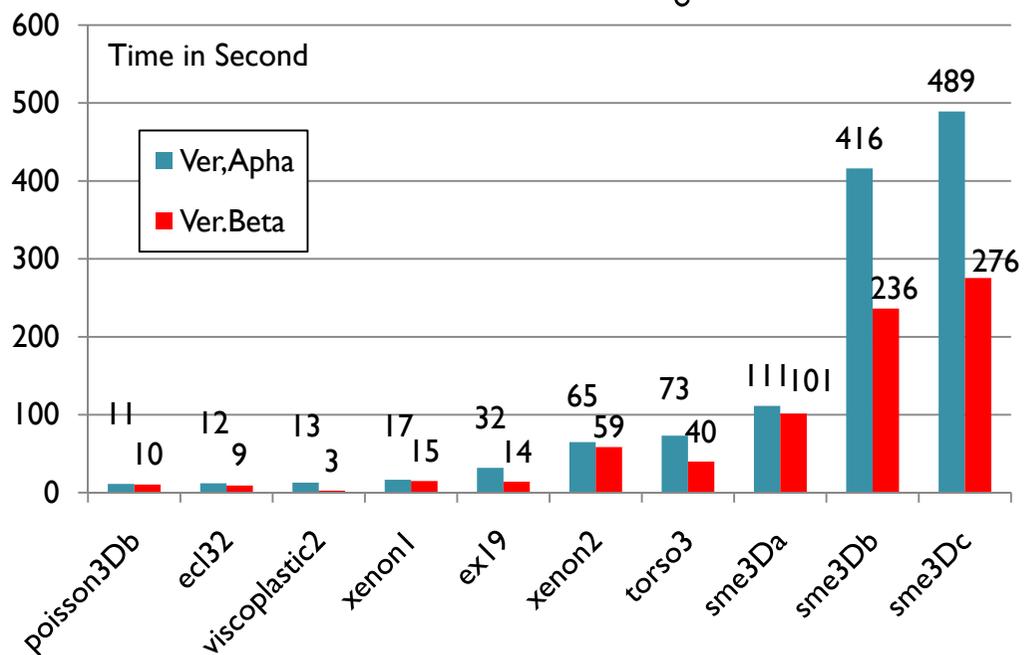
精度ポリシーの効果(OpenATI_LINEAR SOLVE)



torso1 は双方で収束せず。
 α版は以下の行列で相対残差の要求精度を満たせず

- torso3 : 2.28E-08
- memplus : 1.92E-08
- airfoil_2d : 5.17E-05
- viscoplastic2 : 3.67E-05
- wang4 : 4.35E-08

平均速度向上
 1.69倍
 最大速度向上:
 5.29倍 (airfoil_2d)



速度向上は、
 β版で提供される
 高精度再直交化
 ルーチン
 (DGKS 型G-S)
 による収束性の
 向上が理由。

まとめ

- 汎用自動チューニングインターフェース
OpenATLibの提供

以下の新規開発機能

1. ソルバレベルのアルゴリズム選択API
 - OpenATI_EIGENSOLVE と OpenATI_LINEAR SOLVE
 2. 数値計算ポリシー設定API
- 先進的な自動チューニング機能の提供
 1. 数値計算ポリシー機能
 - 計算精度ポリシー:
同じ要求精度を達成する α 版の実行時間と比べ：
 - 約1.6倍 ~2.4倍の速度向上
 - 偽収束（解法中で収束と判定するが残差の誤差を満たさない）の回避
 - 精度要求を達成する手動チューニングの自動化による開発コスト削減
 2. 実行時の直交化アルゴリズムの選択
 - 解の精度向上と、速い収束による速度向上を達成

Xabclib はフリーウェアです！

- Xabclib (ver. Beta) のソースコードは LGPL で、PC クラスタコンソーシアムから配布中です。

- 以下を参照してください。

<http://www.pccluster.org/ja/score7.html>